



Karmaveer Bhaurao Patil University, Satara

**Syllabi for
B. Sc. III Data Science**

**Under
Faculty of Science and Technology
(As per NEP 2020)**

With effect from Academic Year 2026-2027

Syllabus for Bachelor of Science (B. Sc.) Part – III Data Science

Sem (Level)	Courses			OE	VSC/ SEC	AEC/ VEC/ IKS	OJT/ FP/ CEP/ CC/ RP	Total Credits	Degree/ Cum. Cr. MEME
	Course I	Course II	Course III						
	Major Mandatory	Major Elective	-	-					
Sem V (5.5)	Major IX(2) Major X(2) Major XI(2) Major P V (2)	Major XII (Elect.)(2) Major P (Elect.) VI(2)	-	-	VSC III (2) (Major Specific)	AEC III (2) (English)	OJT(4) CEP(2)	22	UG Degree: 132
Sem VI(5.5)	Major XIII(2) Major XIV(2) Major XV(2) Major P VII (2)	Major XVI (Elect.)(2) Major P (Elect.) VIII(2)	-	-	VSC IV (P) (2) (Major Specific) SEC III (2)(P)	AEC IV (2) (English)	FP(2) CC(2)	22	
Credits	16	08	-	-	06	04	10	44	
Total	48	24	12	08	14	16	10	132	Exit Option

COURSE TITLES

1) Semester - V

Level	Semester	Course	Course Code	Course Titles	No. of Lectures/ Practical Per Week	Credits
5.5	V	Major IX	BDST 351	Cloud Computing	2	2
		Major X	BDST 352	Data Structures and Algorithms Using Python	2	2
		Major XI	BDST 353	Theory of Testing of Hypothesis	2	2
		Major XII (Elective)	E-I BDST 354	Deep Learning	2	2
			E-II BDST 355	Cyber Security		
		Major (P) V	BDSP 356	Practical V	1	2
		Major (P) VI (Elective)	BDSP 357	Practical VI	1	2
		VSC III	BDSPVSC III	Data Analytics using Tableau	1	2
		OJT	BDSPOJT	On Job Training	-	4
		CEP	BDSPCEP	Community Engagement Program	-	2
		AEC III	BDSTAEC-III	English (AEC III)	2	2
		Total				22

2) Semester - VI

Level	Semester	Course	Course Code	Course Titles	No. of Lectures/ Practical Per Week	Credits
5.5	VI	Major XIII	BDST 361	Data Mining	2	2
		Major XIV	BDST 362	Time Series Analysis	2	2
		Major XV	BDST 363	Git Hub	2	2
		Major XVI (Elective)	E-I BDST 364	Data Engineering	2	2
			E-II BDST 365	Data Governance Methodology		
		Major (P) VII	BDSP 366	Practical VII	1	2
		Major (P) VIII (Elective)	BDSP 367	Practical VIII	1	2
		VSC IV	BDSPVSC IV	Data Visualization Tool using Power BI	1	2
		SEC III	BDSPSEC III	Healthcare Analytics	1	2
		FP	BDSPFP I	Field Project in Data Science	1	2
		CC	BDSPCC I	Co-curricular Course	1	2
		AEC IV	AEC IV	English	1	2
		Total				22

EVALUATION STRUCTURE:

➤ B.Sc. III NEP 2.0

Semester V

Vertical	Course Category	Course Code	CCE	ESE	Total Marks	Credits
Major	T	BDST 351	20	30	50	02
	T	BDST 352	20	30	50	02
	T	BDST 353	20	30	50	02
	P	BDSP 356	-	50	50	02
Major (1 Theory Paper Out of 2)	T	BDST 354	20	30	50	02
	T	BDST 355	20	30	50	02
	P	BDSP 357	-	50	50	02
AEC	T	AEC III	20	30	50	02
VSC	P	BDSPVSC III	-	50	50	02
CEP	P	BDSPCEP I	-	50	50	02
OJT	P	BDSPOJT I	-	100	100	04
Total					550	22

Semester VI

Vertical	Course Category	Course Code	CCE	ESE	Total Marks	Credits
Major	T	BDST 361	20	30	50	02
	T	BDST 362	20	30	50	02
	T	BDST 363	20	30	50	02
	P	BDSP 366	-	50	50	02
Major (Elective) (1 Theory Papers Out of 2)	T	BDST 364	20	30	50	02
	T	BDST 365	20	30	50	02
	P	BDSP 367	-	50	50	02
AEC	T	AEC IV	20	30	50	02
VSC	P	BDSPVSC IV	-	50	50	02
SEC	P	BDSPSEC III	-	50	50	02
CC	P	BDSPCC I	-	50	50	02
FP	P	BDSPFP I	-	50	50	02
Total					550	22

Abbreviations: DSC: Discipline Specific Course; DSE: Discipline Specific Elective; CCE: Continuous Comprehensive Examination; AEC: Ability Enhancement Course; VSC: Vocational Skill Course; SEC: Skill Enhancement Course; OJT: On Job Training; FP: Field Project; CEP: Community Engagement Program; CC: Co-curricular Course; T: Theory; P: Practical

INTERNAL EVALUATION STRUCTURE

- Internal evaluation of theory papers shall carry a weightage of 40%.
- For the internal evaluation of theory papers, departments have the flexibility to adopt a variety of assessment methods, including class tests, Home Assignments, Short and Long Answer question test, Seminars, Open Book Test, Quiz, case studies, Book Review, fieldwork, Industry Visit for CCE I and CCE II.
- CCE III shall consist of multiple-choice questions.
- Each CCE will carry 10 marks for 2-credit theory papers and 20 marks for 4-credit theory papers, and the best two scores will be considered for final mark entry.
- Internal evaluation of practical papers shall carry a weightage of 40%.
- For practical examinations, internal evaluation will be based on student performance, attendance, journal records and viva.

Bachelor of Science (B. Sc.) Part III

Semester - V

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc. III	Semester - V
	Type: Major IX	Marks: 50
	Credits: 2	From: A. Y. 2026-27
	Name of the Course and Code: Cloud Computing (BDST 351)	
Course Objectives: Students should 1) Understand cloud computing and its key characteristics. 2) Learn the Cloud Deployment Models 3) Compare and contrast public, private, hybrid and multi-cloud deployment models. 4) Understand the architecture of cloud computing systems.		
Course Outcomes: Students will be..... 1. Differentiate cloud service models (IaaS, PaaS, SaaS) and deployment models (Public, Private, Hybrid and Community Cloud). 2. Apply cloud computing concepts and key characteristics to practical scenarios. 3. Analyze the services and offerings of AWS, Microsoft Azure and Google Cloud Platform (GCP). 4. Evaluate and compare cloud service models and cloud platforms to recommend suitable cloud solutions for different application requirements.		
Modules	Title and Content	Hrs.
Module 1:	Cloud Services and Cloud Models 1.1 Introduction to Cloud Computing vs. Cluster vs. Grid Computing 1.2 Cloud Service Models, Introduction to IaaS, PaaS, SaaS, Characteristics and Benefits of Each Model, Security Aspects of Service Models, Applications of IaaS, PaaS, SaaS 1.3 Cloud Deployment Models, Public Cloud, Private Cloud, Hybrid Cloud	08
Module 2:	Cloud Architecture 2.1 Cloud Computing Architecture. 2.2 Community Cloud. 2.3 Amazon Web Services (AWS), Amazon EC2 (Elastic Compute Cloud), Amazon ECS (Elastic Container Service), AWS Lambda for Serverless Computing.	08

	2.4 Scalability in Cloud, Service-Oriented Architecture (SOA), Auto Scaling, Load Balancing, Definition and Basic Concepts, Cloud APIs (RESTful APIs), Relating SOA to Cloud Computing.	
Module 3:	Multi Core Architecture 3.1 Introduction to Multi-Core Architecture. 3.2 Parallel Programming in Cloud. 3.3 Cloud Migration, Migration Strategies, Migration Processes. 3.4 Cloud Bursting, Data Migration in Cloud.	07
Module 4:	Cloud Security 4.1 Working with Data Sources. 4.2 Data Ingestion Techniques, From CSV, From Excel, From SQL Databases. 4.3 Data Preprocessing and Cleaning in Cloud. 4.4 Database Integration Using Python, Creating and Searching Tables, Reading and Storing Configuration Information, Programming with Database Connections.	07
Reference Books:- 1) Rajkumar Buyya, Cloud Computing: Principles and Paradigms, New Delhi: Wiley, 2013. 2) Vijay Madisetti and Arshdeep Bahga, Cloud Computing: A Hands-on Approach, Hyderabad: Orient Blackswan, 2014. 3) Kailash Jayaswal, Cloud Computing Black Book, New Delhi: Dreamtech Press, 2014. 4) Srinivasan, Cloud Computing: A Practical Approach for Learning and Implementation, New Delhi: Pearson Education India, 2014. 5) Rajiv Misra, Cloud and Distributed Computing: Algorithms and Systems, New Delhi: Wiley, 2020. 6) Rajkumar Buyya, Mastering Cloud Computing, New Delhi: McGraw Hill, 2024. 7) M. Sudheep Elayidom, Cloud Computing & Big Data: From the Basics to Practical Use Cases, New Delhi: Cengage Learning India Pvt. Ltd., 2024.		
Evaluation Pattern:		
Total Marks: 50		
Internal Continuous Evaluation (20 Marks): • CCE - I: 10 Marks: Activity Based • CCE - II: 10 Marks: Activity Based • CCE – III: 10 Marks: MCQ Based Note: Best two scores will be considered for final mark entry.		End Semester Examination (30 Marks): • Question -1: Solve the following questions (Five questions of 2 Marks) • Question -2: Attempt any two questions (Three questions of 10 Marks) • Question -3: Attempt any four questions (Six questions of 5 Marks) Note: Conversion of 50 marks of ESE evaluation to 30Marks

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology		
	Yashavantrao Chavan Institute of Science, Satara		
	Board of Studies in Computer Science		
	Programme: B.Sc.	Semester - V	
	Type: Major X	Marks: 50	
	Credits: 2	From: A. Y. 2026-27	
Name of the Course and Code: Data Structures and Algorithms using Python (BDST 352)			
Course Objectives: Students should ... 1) Understand the fundamental concepts of data structures and their importance in solving computational problems. 2) Learn and apply algorithmic strategies for searching, sorting and managing data efficiently using python. 3) Analyze the performance and complexity (time and space) of algorithms using big-o notation. 4) Implement linear and non-linear data structures (such as stacks, queues, linked lists, trees and graphs) in python.			
Course Outcomes: Students will be.... 1) Describe and distinguish between various data structures and their use-cases. 2) Implement linear data structures such as arrays, stacks, queues and linked lists using python. 3) Apply non-linear data structures such as trees and graphs to solve hierarchical and network-based problems. 4) Evaluate and compare the efficiency of sorting and searching algorithms.			
Modules	Title and Contents		Hrs
Module 1:	Foundations of Algorithm and Data Structures 1.1 Fundamentals of Data Structures, Need of Data Structure, Definition of Data Structure, Types of Data Structures 1.2 Algorithm Complexity Analysis, Complexity of Algorithms, Space Complexity, Time Complexity. 1.3 Asymptotic Notation and Efficiency Measures, Asymptotic Notations, Big-O, Theta and Omega, Standard Measures of Algorithmic Efficiency		08
Module 2:	Non-Linear Data Structures 2.1 Trees and Binary Trees, Trees, Binary Trees (Traversal: Inorder, Preorder, Postorder), Binary Search Trees (BST), Tree Operations (Insert, Delete, Search). 2.2 Graphs and Representations, Graphs, Graph Representation (Adjacency List & Matrix). 2.3 Graph Traversals and Pathfinding, Graph Traversals (BFS, DFS), Shortest Path Algorithms (Dijkstra, Bellman-Ford).		08

Module 3:	Linked List 3.1 Introduction to Linked List, List as a Data Structure, Differences with Array, Dynamic Implementation of Linked List, Internal and External Pointers 3.2 Types and Operations on Linked List, Singly Linked List, Doubly Linked List, Circular Linked List, Operations – Create, Traverse, Insert, Delete, Search, Sort, Reverse, Concatenate, Merge. 3.3 Applications and Complexity, Time Complexity of Operations	07
Module 4:	Hashing, Dictionaries, Searching and Sorting Algorithms 4.1 Hashing and Hash Tables, Hashing Concepts, Hash Functions, Collision Resolution Techniques (Chaining, Open Addressing) 4.2 Dictionaries in Python, Dictionaries, Dictionary Operations (Insert, Update, Delete, Search), Applications of Hash Tables, Searching and Sorting Algorithms 4.3 Searching, Linear Search, Binary Search, Sorting – Bubble Sort, Selection Sort, Insertion Sort, Merge Sort, Quick Sort	07
Reference Books:- 1) M. T. Goodrich, R. Tamassia, and D. M. Mount, Data Structures and Algorithms in C++, Hoboken: John Wiley & Sons Inc., 2009. 2) N. Karumanchi, Data Structures and Algorithms for GATE, Hyderabad: CareerMonk Publications, 2011. 3) N. Karumanchi, Data Structure and Algorithmic Thinking with Python, Hyderabad: CareerMonk Publications, 2015. 4) T. Mailund, The Joys of Hashing, New York: Apress, 2019. 5) A. Chatterjee, Linked List Problems: Algorithms that Make You a National Programmer, Chennai: Notion Press, 2022. 6) W. A. Martin and D. N. Ness, Optimizing Binary Trees Grown with a Sorting Algorithm, Charleston: Legare Street Press, 2022. 7) N. Karumanchi, Data Structures and Algorithms Made Easy: Data Structures and Algorithmic Puzzles, Hyderabad: CareerMonk Publications, 2023.		
Evaluation Pattern:		
Total Marks: 50		
Internal Continuous Evaluation (20 Marks): • CCE - I : 10 Marks: Activity Based • CCE - II: 10 Marks: Activity Based • CCE – III: 10 Marks: MCQ Based Note: Best two scores will be considered for final mark entry.		End Semester Examination (30 Marks): • Question -1: Solve the following questions (Five questions of 2 Marks) • Question -2: Attempt any two questions (Three questions of 10 Marks) • Question -3: Attempt any four questions (Five questions of 5 Marks) Note: Conversion of 50 marks of ESE evaluation to 30Marks

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester - V
	Type: Major XI	Marks: 50
	Credits: 2	From: A. Y. 2026-27
Name of the Course and Code: Theory of Testing of Hypothesis (BDST 353)		
Course Objectives: Students should be... 1) Understand the fundamental concepts of hypothesis testing and the difference between parametric and non-parametric methods. 2) Learn how to apply appropriate statistical tests based on data type, distribution and sample size. 3) Develop practical skills for performing hypothesis tests on real-world data. 4) Interpret and communicate statistical test results effectively in the context of scientific or business decision-making.		
Course Outcomes: Students will be... 1) Identify and apply suitable parametric or non-parametric tests for different research questions and data types. 2) Perform hypothesis testing using tools such as R, Python or Excel. 3) Evaluate the assumptions behind statistical tests and choose alternative methods when assumptions are violated. 4) Interpret p-values, confidence intervals and test statistics to support data-driven conclusions and decisions.		
Module	Title and Contents	Hrs.
Module 1:	Fundamentals of Hypothesis Testing 1.1 Introduction to Inferential Statistics, Population and Sample 1.2 Sampling Techniques, Sampling Distribution, Central Limit Theorem 1.3 Concept of Hypothesis Testing, Null Hypothesis (H_0), Alternative Hypothesis (H_1) 1.4 Types of Errors (Type I, Type II), Level of Significance, Power of Test, Steps in Hypothesis Testing	8
Module 2:	Parametric Tests 2.1 Assumptions of Parametric Tests, Z-Test (One-sample, Two-sample) 2.2 T-Test (One-sample, Independent, Paired t-test) 2.3 Confidence Intervals, Relationship with Hypothesis Testing 2.4 P-value Approach, Critical Value Approach, Effect Size (Cohen's d)	8
Module 3:	Non-Parametric and Categorical Tests 3.1 Introduction to Non-Parametric Tests, Need and Applications 3.2 Chi-Square Test (Test of Independence, Goodness of Fit)	8

	3.3 Mann-Whitney U Test, Wilcoxon Signed Rank Test 3.4 Kruskal-Wallis Test, Sign Test, Median Test, Parametric vs Non-Parametric Tests	
Module 4:	Advanced Hypothesis Testing and Applications 4.1 ANOVA (One-way ANOVA, Two-way ANOVA) 4.2 Multiple Hypothesis Testing 4.3 Bonferroni Correction, False Discovery Rate (FDR) 4.4 Experimental Design, Bayesian Hypothesis Testing, Real-world Applications	6
Reference Books:- 1) Ashish Kumar and Gary Dougan, Learning Predictive Analytics with Python (Birmingham, UK: Packt Publishing, 2015). 2) V. Veeraraghavan, Textbook of Parametric and Nonparametric Statistics (New Delhi: Sage Publications India, 2016). 3) Scott Hartshorn, Hypothesis Testing: A Visual Introduction to Statistical Significance (Independently Published, 2017). 4) S. C. Gupta and V. K. Kapoor, Applied Statistics (New Delhi: Sultan Chand & Sons, 2018). 5) K. Suresh and Usman Ahmed, Hands-On Exploratory Data Analysis with Python (Birmingham, UK: Packt Publishing, 2020). 6) S. K. Sharma, Testing of Hypothesis and Applied Statistics (New Delhi: New Academic Science, 2023). 7) Gareth James, Daniela Witten, and Trevor Hastie, An Introduction to Statistical Learning: With Applications in Python (Springer Nature, 2023).		
Evaluation Pattern:		
Total Marks: 50		
Internal Continuous Evaluation (20 Marks): • CCE - I: 10 Marks: Activity Based • CCE - II: 10 Marks: Activity Based • CCE – III: 10 Marks: MCQ Based Note: Best two scores will be considered for final mark entry.		End Semester Examination (30 Marks): • Question -1: Solve the following questions (Five questions of 2 Marks) • Question -2: Attempt any two questions (Three questions of 10 Marks) • Question -3: Attempt any four questions (Five questions of 5 Marks) Note: Conversion of 50 marks of ESE evaluation to 30Marks

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester - V
	Type: Major P V	Marks: 50
	Credits: 2	From: A. Y. 2026-27
Name of the Course and Code: Practical V BDSP 356		
Course Objectives: Students should...		
1) Understand cloud computing and its key characteristics. 2) Differentiate between infrastructure as a service (IAAS), platform as a service (PAAS), and software as a service (SAAS). 3) Understand the memory requirement for various data structures. 4) Operate on the various structured data.		
Course Outcomes: Students will be...		
1) Demonstrate proficiency in building various Tableau visualizations, from basic charts to advanced analytics. 2) Design and publish interactive dashboards tailored to business needs. 3) Efficiently handle large datasets using Tableau's data preparation and performance optimization features. 4) Apply Tableau for real-world scenarios, leveraging advanced features like LOD calculations, mapping, and integrations.		
List of Experiments		Hrs
1. Create an Amazon AWS Account and create Amazon S3 Bucket. 2. Implementation VCP and creation Amazon EC2 Instance (Window) 3. Amazon Elastic Load Balancer (ELB) 4. Amazon RDS & DynamoDB 5. Deploying machine learning models into production as time-consuming and error-prone. 6. Perform Python program to search for an element in a list using linear search. and binary search on a sorted list. 7. Implement binary search using recursion in Python. 8. Perform a Python program to implement bubble sort, selection sort, insertion sort and quick sort 9. One-Sample t-test, Independent Two-Sample t-test and Paired Sample t-test 10. Z-test for Mean and Z-test for Proportion 11. ANOVA (Analysis of Variance)		60

12. Pearson's Correlation Coefficient, Simple Linear Regression and Multiple Linear Regression 13. Levene's Test for Equality of Variances 14. Chi-Square Test of Independence 15. Mann–Whitney U Test	
Reference Books: 1) Michael T. Goodrich, Roberto Tamassia, and David M. Mount, Data Structures and Algorithms in C++, Hoboken: John Wiley & Sons Inc., 2009. 2) Narasimha Karumanchi, Data Structures and Algorithms for GATE, Hyderabad: CareerMonk Publications, 2011. 3) Kailash Jayaswal, Cloud Computing Black Book, New Delhi: Dreamtech Press, 2014. 4) Stephen Armstrong, Codeless Data Structures and Algorithms: Learn DSA Without Writing a Single Line of Code, New York: Apress, 2020. 5) W. A. Martin and D. N. Ness, Optimizing Binary Trees Grown with a Sorting Algorithm , Charleston: Legare Street Press, 2022. 6) Rajkumar Buyya, Mastering Cloud Computing, 2nd ed., New Delhi: McGraw Hill, 2024. 7) M. Sudheep Elayidom, Cloud Computing & Big Data: From the Basics to Practical Use Cases, New Delhi: Cengage Learning India Pvt. Ltd., 2024.	
Evaluation Pattern:	
Total Marks: 50	
Journal, Students' Performance Viva, Project (20 Marks): • Journal: 10 Marks • Students' Performance: 05 Marks • Viva: 5 Marks	Practical Exam Paper (30 Marks): • Section I: Attempt any two questions (Three questions of 10 Marks) • Section II: Attempt any two questions (Two questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester - V
	Type : Major XII (Elective)	Marks: 50
	Credits : 2	From: A. Y. 2026-27
Name of the Course and Code: Deep Learning (BDST 354)		
Course Objectives: Students should ... 1) Learn how neurons, layers, weights, and biases work 2) Grasp knowledge about forward and backward propagation 3) Understand loss functions and gradient descent 4) Explore techniques like regularization, dropout, and batch normalization		
Course Outcomes: Students will be... 1) Build and train neural networks from scratch or using libraries 2) Tune hyperparameters for better performance 3) Apply deep learning to real-world datasets 4) Evaluate models using metrics (accuracy, precision, recall, etc.)		
Module	Title and Contents	Hrs.
Module 1:	Fundamentals of Neural Network 1.1 Biological neuron, Mc-Culloch Pitts Neuron, 1.2 Perceptron, Perceptron Learning, Delta learning 1.3 Multilayer Perceptron: Linearly separable, linearly non-separable classes, 1.4 Deep Networks: Fundamentals, Brief History, Supervised vs. Unsupervised Learning	6
Module 2:	Introduction to Deep learning 2.1 Introduction to Deep Feed forward Networks, 2.2 Gradient Based Learning, 2.3 Hidden Units Architecture-Design, 2.4 Epcoh, Underfitting and Overfitting	6
Module 3:	Regularization and Optimization Techniques 3.1 Introduction to Non-Parametric Tests, Need and Applications Chi-Square Test (Test of Independence, Goodness of Fit) 3.2 Mann-Whitney U Test, Wilcoxon Signed Rank Test	8

	3.3 Kruskal-Wallis Test, Sign Test, Median Test, Parametric vs Non-Parametric 3.4 Tests Regularization: Need of Regularization, L2 Regularization, L1 Regularization, Early Stopping and Dropout, 3.5 Optimization: Challenges in NN Optimization, Gradient Descent Approaches, Parameter Initialization Approach, Adaptive Approaches - AdaGrad, RMSProp and Adam, Introduction to Batch Normalization	
Module 4:	Advanced Neural Network Architectures 4.1 Introduction of various CNN Architectures: LeNet, AlexNet, VGG, GoogleNet, and UNet, 4.2 Applications of CNN in Image Classification and Object Detection 4.3 Recurrent and Recursive Nets: Recurrent Neural Networks, 4.4 Bidirectional RNN, Encoder Decoder Architectures	8
Reference Books:- 1) Christopher M. Bishop, Pattern Recognition and Machine Learning (New York: Springer, 2006). 2) Michael A. Nielsen, Neural Networks and Deep Learning, Determination Press, 2015. 3) Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, Cambridge: MIT Press, 2016. 4) François Chollet, Deep Learning with Python, Shelter Island, NY: Manning Publications, 2018. 5) Aurélien Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Sebastopol, CA: O'Reilly Media, 2019. 6) Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola, Dive into Deep Learning, Cambridge: MIT Press, 2021. 7) Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative Adversarial Nets." In Advances in Neural Information Processing Systems 27, 2672–2680. Red Hook, NY: Curran Associates, 2014.		
Evaluation Pattern:		
Total Marks: 50		
Internal Continuous Evaluation (20 Marks): • CCE - I: 10 Marks: Activity Based • CCE - II: 10 Marks: Activity Based • CCE – III: 10 Marks: MCQ Based Note: Best two scores will be considered for final mark entry.		End Semester Examination (30 Marks): • Question -1: Solve the following questions (Five questions of 2 Marks) • Question -2: Attempt any two questions (Three questions of 10 Marks) • Question -3: Attempt any four questions (Five questions of 5 Marks) Note: Conversion of 50 marks of ESE evaluation to 30Marks

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc. III	Semester - V
	Type: Major P VI (Elective)	Marks: 50
	Credits: 2	From: A. Y. 2026-27
	Name of the Course and Code: Practical VI BDSP 357	
Course Objectives: Students will be able to...		
1) Understand the fundamentals of Artificial Neural Networks and deep learning concepts. 2) Study various neural network architectures such as Convolutional Neural Networks, Recurrent Neural Networks, and modern models. 3) Learn training techniques including forward propagation, backpropagation, and optimization methods. 4) Gain practical knowledge of implementing deep learning models using tools like TensorFlow and PyTorch.		
Course Outcomes: By the end of the course, students will be able to:		
1) Design and implement different deep learning architectures for various applications. 2) Analyze and evaluate model performance using appropriate metrics. 3) Apply deep learning techniques to solve real-world problems in areas like computer vision and NLP. 4) Use modern frameworks such as TensorFlow and PyTorch for model development and deployment.		
List of Experiments		Hrs.
1) Study and implementation of a single-layer perceptron and multi-layer neural network (MLP) 2) Study of activation functions (ReLU, Sigmoid, Tanh) 3) Implementation of gradient descent optimization 4) Solving XOR problem using neural networks 5) Dataset loading and preprocessing using PyTorch 6) Implementation of feedforward neural network for classification 7) Implementation of Convolutional Neural Networks for image classification 8) Training CNN on MNIST dataset		60

9) Data augmentation techniques for image datasets 10) Transfer learning using pretrained models 11) Implementation of Recurrent Neural Networks 12) Sequence modeling using LSTM networks 13) Sentiment analysis using deep learning models 14) Study of attention mechanism in deep learning 15) Comparison of optimization algorithms (SGD, Adam, RMSProp) 16) Implementation of dropout and regularization techniques	
Reference Books:- 1) Christopher M. Bishop, Pattern Recognition and Machine Learning (New York: Springer, 2006). 2) Michael A. Nielsen, Neural Networks and Deep Learning (Determination Press, 2015). 3) Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning (Cambridge: MIT Press, 2016). 4) François Chollet, Deep Learning with Python (Shelter Island, NY: Manning Publications, 2018). 5) Aurélien Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (Sebastopol, CA: O'Reilly Media, 2019). 6) Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola, Dive into Deep Learning (Cambridge: MIT Press, 2021). 7) Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative Adversarial Nets." In Advances in Neural Information Processing Systems 27, 2672–2680. Red Hook, NY: Curran Associates, 2014.	
Evaluation Pattern:	
Total Marks: 50	
Journal, Students' Performance Viva, Project (20 Marks): • Journal: 10 Marks • Students' Performance: 05 Marks - Viva: 5 Marks	Practical Exam Paper (30 Marks): • Section I: Attempt any two questions (Three questions of 10 Marks) • Section II: Attempt any two questions (Two questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology		
	Yashavantrao Chavan Institute of Science, Satara		
	Board of Studies in Statistics (Data Science)		
	Programme: B.Sc.	Semester - V	
	Type : Major XII (Elective)	Marks: 50	
	Credits : 2	From: A. Y. 2026-27	
Name of the Course and Code: Cyber Security (BDST 355)			
Course Objectives: Students should...			
1) Understand the basics of computer, network and information security. 2) Study operating system security and malwares. 3) Acquaint with security issues in internet protocols. 4) Analyze the system for vulnerabilities.			
Course Outcomes: Students will be			
1) Use cryptographic techniques in secure application development. 2) Apply methods for authentication, access control, intrusion detection and prevention. 3) Apply the scientific method for security assessment. 4) Illustrate computer forensics knowledge.			
Module	Title and Contents		Hrs
Module 1:	Basic concepts of Cyber Security 1.1 Introduction to Cyber Security Concepts, Elements of Information Security, Security Policies and Techniques, 1.2 Security Process and Categories, Steps in Information Security, Categories of Security Measures. 1.3 Network Security Basics, Operational Model of Network Security, Basic Terminologies in Network Security. 1.4 Threats and Vulnerability, Definition and Types of Threats, Identifying System Vulnerabilities 1.5 Security vs Privacy, Understanding the Differences		8
Module 2:	Data Encryption Techniques And Standards 2.1 Introduction to Encryption, Encryption Concepts and Need, Encryption Techniques, Symmetric vs Asymmetric Encryption, Basics of Cryptography 2.3 Classical Encryption Methods, Substitution and Transposition Ciphers, Steganography – Applications and Limitations		6

	2.4 Modern Encryption Algorithms, Block Ciphers and Modes of Operation, Feistel Cipher Structure 2.5 Standard Encryption Algorithms, Data Encryption Standard (DES), Triple DES, Weak Keys in DES, Advanced Encryption Standard (AES)	
Module 3:	Public Key and Management 3.1 Public Key Cryptography Overview, RSA Algorithm – Working, Key Length, Security 3.2 Key Exchange and Management, Key Distribution Techniques, Diffie-Hellman Key Exchange 3.3 Elliptic Curve Cryptography, ECC Arithmetic and Cryptography, ECC Security and Applications 3.4 Authentication and Integrity, Message Digest and Hashing, Kerberos and X.509 Authentication Service 3.5 Digital Signatures, Implementation and Algorithms, DSS and Authentication Protocols.	8
Module 4:	Computer Intrusion and Security Mechanisms 4.1 Computer Intrusions 4.2 Firewall Fundamentals, Characteristics and Types of Firewalls, Benefits and Limitations of Firewalls, Firewall Architecture 4.3 Trusted Systems and Access Control Models 4.4 Intrusion Detection Systems (IDS), IDS Concepts, Need, and Methods, Types of IDS and Their Implementation 4.5 Password Management, Password Policies and Protection, Limitations and Security Challenges.	8
Reference Books:- 1) Christopher M. Bishop, Pattern Recognition and Machine Learning, New York: Springer, 2006. 2) Michael A. Nielsen, Neural Networks and Deep Learning, Determination Press, 2015. 3) Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, Cambridge: MIT Press, 2016. 4) François Chollet, Deep Learning with Python, Shelter Island, NY: Manning Publications, 2018. 5) Aurélien Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Sebastopol, CA: O'Reilly Media, 2019. 6) Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola, Dive into Deep Learning, Cambridge: MIT Press, 2021.		

- 7) Stallings, William, and Lawrie Brown. Computer Security: Principles and Practice. 4th ed. Boston: Pearson, 2018.
- 8) Forouzan, Behrouz A. Cryptography and Network Security. 3rd ed. New York: McGraw-Hill Education, 2015.

Evaluation Pattern:

Total Marks: 50

Internal Continuous Evaluation (20 Marks):

- CCE - I: 10 Marks: Activity Based
- CCE - II: 10 Marks: Activity Based
- CCE – III: 10 Marks: MCQ Based

Note: Best two scores will be considered for final mark entry.

End Semester Examination (30 Marks):

- Question -1: Solve the following questions (Five questions of 2 Marks)
- Question -2: Attempt any two questions (Three questions of 10 Marks)
- Question -3: Attempt any four questions (Five questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc. III	Semester - V
	Type: Major P VI (Elective)	Marks: 50
	Credits: 2	From: A. Y. 2026-27
	Name of the Course and Code: Practical VI BDSP 357	
Course Objectives: Students should to...		
1) Understand the basics of computer, network and information security. 2) Study operating system security and malwares. 3) Acquaint with security issues in internet protocols. 4) Analyze the system for vulnerabilities.		
Course Outcomes: students will be able to:		
1) Use cryptographic techniques in secure application development. 2) Apply methods for authentication, access control, intrusion detection and prevention. 3) Evaluate the scientific method for security assessment. 4) Illustrate computer forensics knowledge.		
List of Experiments		Hrs
1) Practical on cyber-attacks. 2) Hands on Network Security. 3) Exercise on Symmetric encryption algorithm. 4) practical on Asymmetric Symmetric encryption algorithm. 5) Exercise on Intrusion detection. 6) practical on Cryptography. 7) Hands on Public encryption algorithm. 8) practical on Digital Signature 9) Exercise on Firewall And Intrusion. 10) Identification of Security Threats and Vulnerabilities 11) Implementation of Classical Encryption Techniques 12) Implementation of Symmetric Encryption using AES 13) Implementation of Content-Based Recommender System 14) Implementation of Collaborative Filtering 15) Implementation of Hybrid Recommendation System		60

Reference Books:-

- 1) Christopher M. Bishop, Pattern Recognition and Machine Learning (New York: Springer, 2006).
- 2) Michael A. Nielsen, Neural Networks and Deep Learning (Determination Press, 2015).
- 3) Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning (Cambridge: MIT Press, 2016).
- 4) François Chollet, Deep Learning with Python (Shelter Island, NY: Manning Publications, 2018).
- 5) Aurélien Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (Sebastopol, CA: O'Reilly Media, 2019).
- 6) Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola, Dive into Deep Learning (Cambridge: MIT Press, 2021).
- 7) Northcutt, Stephen, Lenny Zeltser, Scott Winters, Karen Kent, and Ronald W. Ritchey. Inside Network Perimeter Security. 2nd ed. Indianapolis, IN: Sams Publishing, 2005.
- 8) Scarfone, Karen, and Peter Mell. Guide to Intrusion Detection and Prevention Systems (IDPS). Gaithersburg, MD: National Institute of Standards and Technology, 2007.

Evaluation Pattern:**Total Marks: 50****Journal, Students' Performance Viva, Project (20 Marks):**

- Journal: 10 Marks
- Students' Performance: 05 Marks
- Viva: 5 Marks

Practical Exam Paper (30 Marks):

- Section I: Attempt any two questions (Three questions of 10 Marks)
- Section II: Attempt any two questions (Two questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc. III	Semester - V
	Type : VSC III	Marks: 50
	Credits : 2	From: A. Y. 2026-27
	Name of the Course and Code: Data Analytics using Tableau (BDSPVSC III)	
Course Objectives: Students will be able to...		
1) Equip students with foundational to advanced tableau skills for data visualization and analytics. 2) Enable learners to effectively connect, prepare, and transform data for insights. 3) Teach the creation of dynamic and interactive dashboards and visualizations. 4) Prepare students for real-world applications and tableau certification.		
Course Outcomes: Students will be....		
1) Demonstrate proficiency in building various tableau visualizations, from basic charts to advanced analytics. 2) Design and publish interactive dashboards tailored to business needs. 3) Efficiently handle large datasets using tableau's data preparation and performance optimization features. 4) Apply tableau for real-world scenarios, leveraging advanced features like LOD calculations, mapping, and integrations.		
List of Experiments		Hrs
1) Installing Tableau Desktop and navigate its interface. 2) Automate basic data cleaning and explore metadata in Tableau using data interpreter. 3) Extract data from a live connection and optimize for performance. 4) Create heat maps, tree maps, and scatter plots for advanced visualizations. 5) Combine sheets into a dashboard with interactive filters and actions. 6) Develop calculated fields for custom metrics and KPIs. 7) Implement running totals, moving averages, and percent of totals 8) Apply LOD (Level of Detail) expressions to perform advanced calculations. 9) Perform data blending and data joining from multiple data sources. 10) Design and implement interactive dashboards with drill-down and highlight actions. 11) Use filters (context, top, condition, and quick filters) effectively in dashboards. 12) Create story points to present data insights in a structured narrative format. 13) Implement forecasting and trend lines for predictive analysis. 14) Apply formatting and styling techniques for professional dashboard design. 15) Optimize dashboard performance using performance recording and best practices		60

Reference Books: -

- 1) Ben Jones, Communicating Data with Tableau Designing, Developing, and Delivering Data Visualizations, Sebastopol, CA: O'Reilly Media, 2014.
- 2) Daniel G. Murray, Tableau Your Data: Fast and Easy Visual Analysis with Tableau Software, Hoboken, NJ: Wiley, 2016.
- 3) Ryan Sleeper, Practical Tableau: 100 Tips, Tutorials, and Strategies from a Tableau Zen Master, Sebastopol, CA: O'Reilly Media, 2020.
- 4) Tim Costello and Lori Sinnamon, Mastering Tableau 2021, Birmingham, UK: Packt Publishing, 2021.
- 5) Carl Allchin, Tableau Prep Up & Running, Sebastopol, CA: O'Reilly Media, 2021.
- 6) Joshua N. Milligan, Learning Tableau 2022: Creating Effective Data Visualizations and Dashboards, Birmingham, UK: Packt Publishing, 2022.
- 7) George Peck, Tableau 2022: The Complete Reference, New York: McGraw Hill, 2022.

Evaluation Pattern:**Total Marks: 50****Journal, Students' Performance Viva, Project (20 Marks):**

- Journal: 10 Marks
- Students' Performance: 05 Marks
- Viva: 5 Marks

Practical Exam Paper (30 Marks):

- Section I: Attempt any two questions (Three questions of 10 Marks)
- Section II: Attempt any two questions (Two questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester: V
	Type: OJT	Marks: 100
	Credits: 4	From: A. Y. 2026-27
	Name of the Course & Code: On Job Training (BDSPOJT)	
Course Objectives: 1) Apply advanced Data manipulation techniques and data analysis methods in a professional setting. 2) Gain hands-on experience with statistical software and real-world datasets. 3) Develop problem-solving, communication and reporting skills in a workplace environment.		
Course Outcomes: 1) Students will effectively apply on real-world industry or research problems. 2) Students will use softwares proficiently to manage and analyze data. 3) Students will handle professional reports and presentations.		
Module	Title and Contents	Hrs.
Module 1:	OJT will provide the opportunities for internship with local/regional industries, business organization, health and allied areas, local government, etc. so that students may actively engaged with the employability opportunities. Students will undergo 4 credit work-based learning/OJT/internship.	120
Evaluation Pattern:		
Total Marks: 100		
• Report submission= 20 Marks • Presentation of report=20 Marks • Viva= 20 Marks • Industry Evaluation=40 Marks		

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester: V
	Type: CEP	Marks: 100
	Credits: 4	From: A. Y. 2026-27
	Name of the Course & Code : Community Engagement Programme (BDSPCEP)	
Course Objectives: 1) Promote respect for rural culture, lifestyle and wisdom among students. 2) Learn about the present status of agricultural and development initiatives. 3) Identify and address the root causes of distress and poverty among vulnerable households. 4) Improve learning outcomes by applying classroom knowledge to real-world situations		
Course Outcomes: 1) After completing this course, students will be able to 2) Gain an understanding of rural life, Indian culture, and social realities. 3) Develop empathy and bonds of mutuality with the local community. 4) Appreciate the significant contributions of local communities to Indian society and economy.		
Module Title	Module Content	Hours
Module 1: Orientation	Introduction to CEP, ethics, community overview, safety, and communication training; concept and importance of community engagement; roles and responsibilities of engaged citizens	15
Module 2: Community Study	Village/slum visit, mapping, observation, data collection, household survey and community interactions.	15
Module 3: Scientific Activity	Demonstrations on health, hygiene, water testing, environmental awareness, etc.	30
Module 4: Reflection & Report	Field diary writing, group discussions, report preparation and presentation.	--
Evaluation Pattern:		
Total Marks:50		
Component	Marks	
Attendance and Participation	10	

Field Diary / Logbook	10
Scientific Activity / Demonstration	10
Final Report	10
Viva Voce / Presentation	10
Total	50

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester - VI
	Type : Major XIII	Marks: 50
	Credits : 2	From: A. Y. 2026-27
Name of the Course and Code: Data Mining (BDST 361)		
Course Objectives: Students should...		
1) Understand the fundamentals of data mining and its applications. 2) Learn Python tools for data handling, preprocessing, and visualization. 3) Comprehend the Data Mining environments and applications. 4) Gain practical experience with real-world datasets.		
Course Outcomes: Students will be.....		
1) Explain the fundamental concepts, applications, and significance of data mining in solving real-world problems. 2) Apply Python programming and libraries such as NumPy, Pandas, and Matplotlib for data loading, preprocessing, exploration, and analysis. 3) Perform data preprocessing techniques, including handling missing values, data transformation, and visualization, to prepare datasets for analysis. 4) Implement classification algorithms, including Decision Tree, Naïve Bayes, and k-Nearest Neighbors (KNN), to classify data and develop predictive models.		
Module	Title and Contents	Hrs.
Module 1:	Introduction to Data Mining 1.1 Overview of Data Mining and Knowledge Discovery 1.2 Data Mining Functionalities and Applications 1.3 Data Mining Process and Architecture 1.4 Data Types, Data Quality, and Challenges 1.5 Python for Data Mining: Environment Setup, Jupyter, Pandas, NumPy Overview	08
Module 2:	Classification Techniques 2.1 Overview of Supervised Learning 2.2 Decision Tree, Naïve Bayes, K-Nearest Neighbor (KNN) 2.4 Support Vector Machines (SVM) and Logistic Regression 2.4 Model Evaluation: Accuracy, Precision, Recall, F1-score 2.5 Implementation of Classification Algorithms in Python (scikit-learn)	07

Module 3:	Clustering and Association Analysis 3.1 Introduction to Clustering – K-Means, Hierarchical, DBSCAN 3.2 Cluster Evaluation and Visualization 3.4 Association Rule Mining – Apriori and FP-Growth 3.4 Market Basket Analysis and Applications 3.5 Implementation of Clustering and Association Algorithms in Python	07
Module 4:	Advanced Topics and Applications 4.1 Text Mining and Sentiment Analysis Basics 4.2 Introduction to Web Mining and Social Media Mining 4.3 Case Studies in Healthcare, E-commerce, and Finance 4.4 Mini Project: Data Mining Workflow on a Real Dataset	08
Reference Books:- 1) Han, J., Kamber, M., & Pei, J. <i>Data Mining: Concepts and Techniques</i>. San Francisco: Morgan Kaufmann, 2011. 2) McKinney, W. <i>Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython</i>. Sebastopol: O'Reilly Media, 2017. 3) Raschka, S. <i>Python Machine Learning</i>. Birmingham: Packt Publishing, 2015. 4) Tan, P.-N., Steinbach, M., & Kumar, V. <i>Introduction to Data Mining</i>. Boston: Addison-Wesley, 2005 5) Patel, R., & Mehta, R. <i>Data Mining Using Python</i>. New Delhi: BPB Publications, 2021. 6) Leskovec, J., Rajaraman, A., & Ullman, J. D. <i>Mining of Massive Datasets Using Python</i>. Cambridge: Cambridge University Press, 2020.		
Evaluation Pattern:		
Total Marks: 50		
Internal Continuous Evaluation (20 Marks): • CCE - I: 10 Marks: Activity Based • CCE - II: 10 Marks: Activity Based • CCE – III: 10 Marks: MCQ Based Note: Best two scores will be considered for final mark entry.	End Semester Examination (30 Marks): • Question -1: Solve the following questions (Five questions of 2 Marks) • Question -2: Attempt any two questions (Three questions of 10 Marks) • Question -3: Attempt any four questions (Five questions of 5 Marks)	

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc. III	Semester - VI
	Type: Major XIV	Marks: 50
	Credits: 2	From: A. Y. 2026-27
Name of the Course and Code: Time Series Analysis (BDST 362)		
Course Objectives: Students should ... 1) Understand fundamental concepts of time series data and components. 2) Equip students with techniques for modeling, analyzing, and forecasting time-dependent data. 3) Develop understanding of time series models such as Arima, exponential smoothing, and others. 4) Provide hands-on experience using software tools.		
Course Outcomes: Students will be ... 1) Decompose components of time series data. 2) Apply and interpret various time series models including exponential smoothing. 3) Use statistical software to perform time series analysis and generate forecasts. 4) Evaluate the accuracy of forecasting models using performance metrics such as mse, rmse		
Module	Title and Contents	Hrs.
Module 1:	Time Series Analysis 1.1 Fundamentals of Time Series, Definition and basic terminology, Characteristics of time series data, Differences between time series and cross-sectional data 1.2 Components of Time Series, Trend component, Seasonal component, Cyclical component, Irregular (random) component.	08
Module 2:	Applications of Time Series Analysis 2.1 Applications of Time Series Analysis, Business and economics, Weather and environment, Health and social sciences 2.2 Working with Time Series Data using Python, Indexing and parsing dates, Loading and visualizing time series with Pandas, Time-based indexing and slicing 2.3 Time Series Decomposition, Additive vs multiplicative models, Seasonal decomposition using moving averages (STL)	07
Module 3:	Preprocessing and Smoothing Techniques 3.1 Data Cleaning and Preprocessing, Handling missing data, Identifying and treating outliers	

	3.2 Smoothing Techniques, Simple moving average, Weighted moving average, Exponential smoothing	07
Module 4:	Time Series Analysis for Data Science 4.1 Real-world Applications and Case Studies, Stock market forecasting, Weather prediction, Demand and sales forecasting data, 4.2 Python Libraries for Time Series, pandas, stats models, prophet, scikit-learn 4.3 TensorFlow, keras, seaborn, matplotlib Libraries.	08
Reference Books:- 1) James D. Hamilton, Time Series Analysis, Princeton: Princeton University Press, 1994. 2) Paul S. P. Cowpertwait and Andrew V. Metcalfe, Introductory Time Series with R, New York: Springer, 2009. 3) George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung, Time Series Analysis: Forecasting and Control, Hoboken: Wiley, 2015. 4) Douglas C. Montgomery, Cheryl L. Jennings, and Murat Kulahci, Introduction to Time Series Analysis and Forecasting, Hoboken: Wiley, 2015. 5) Peter J. Brockwell and Richard A. Davis, Introduction to Time Series and Forecasting, New York: Springer, 2016. 6) Robert H. Shumway and David S. Stoffer, Time Series Analysis and Its Applications: With R Examples, New York: Springer, 2017. 7) Rob J. Hyndman and George Athanasopoulos, Forecasting: Principles and Practice, Melbourne: OTexts, 2018.		
Evaluation Pattern:		
Total Marks: 50		
Internal Continuous Evaluation (20 Marks): • CCE - I: 10 Marks: Activity Based • CCE - II: 10 Marks: Activity Based • CCE – III: 10 Marks: MCQ Based Note: Best two scores will be considered for final mark entry.		End Semester Examination (30 Marks): • Question -1: Solve the following questions (Five questions of 2 Marks) • Question -2: Attempt any two questions (Three questions of 10 Marks) • Question -3: Attempt any four questions (Five questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology		
	Yashavantrao Chavan Institute of Science, Satara		
	Board of Studies in Statistics (Data Science)		
	Programme: B.Sc.	Semester - VI	
	Type : Major XV	Marks: 50	
	Credits : 2	From: A. Y. 2026-27	
Name of the Course and Code: Git Hub (BDST 363)			
Course Objectives: Students should...			
1) Understand the fundamentals of git and github. 2) Apply branching, merging, and collaboration techniques to work effectively on projects with teams. 3) Implement advanced git features such as stashing, rebasing, cherry-picking. 4) Learn project management, and continuous integration/deployment (ci/cd) workflows.			
Course Outcomes: Students will be			
1) Demonstrate proficiency in using git commands for version control and repository management. 2) Collaborate effectively on projects using github features such as pull requests, forks, and issues. 3) Resolve conflicts, manage branches, and automate workflows using git and github actions. 4) Develop and maintain software projects using github best practices.			
Module	Title and Contents		Hrs
Module 1:	Fundamental Concepts of GitHub 1.1 Version Control Systems, Centralized Version Control Systems, Distributed Version Control Systems 1.2 Introduction to Git, Features and Benefits of Git, Installing Git and Configuring User Information 1.3 Basic Git Commands, git init, git clone, git status, git add, git commit, git log 1.4 Understanding Git Workflow, Creating a Repository, Cloning, Pushing, and Pulling, GitHub v s. GitLab vs. Bitbucket		08
Module 2:	GitHub Operations 2.1 Branches in Git, Creating, Switching, and Deleting Branches (git branch, git checkout, git switch) 2.2 Merging Branches, git merge, git rebase 2.3 Resolving Merge Conflicts 2.4 Collaboration on GitHub, Forks, Pull Requests, and Code Reviews, Managing Permissions and Issues in GitHub, GitHub Actions and Workflow Automation		07

Module 3:	Techniques in Github 3.1 Stashing and Cleaning, git stash, git clean 3.2 Undoing Changes, git reset, git revert, git checkout 3.3 Configuration Files, .gitignore, .gitattributes 3.4 Rebasing Techniques, git rebase, git rebase -i (Interactive Rebase), Working with Submodules, Cherry-Picking Commits (git cherry-pick), Git Hooks for Automation	08
Module 4:	GitHub for DevOps and Project Management 4.1 Using GitHub for Open Source Projects 4.2 Managing Large Repositories, Git LFS (Large File Storage) 4.3 Agile Development with GitHub, GitHub Issues, Project Boards 4.4 CI/CD with GitHub, Continuous Integration (CI), Continuous Deployment (CD) using GitHub Actions, Security Best Practices in Git and GitHub, Monitoring and Auditing Changes	07
Reference Books:- 1) Mike McQuaid, Git in Practice: Includes 66 Techniques (New York: Manning, 2014). 2) Phil Haack and Sarah Guthals, GitHub for Dummies (New York: Wiley, 2019). 3) Jawwad Ahmad, Cédric Belanger, and Raywenderlich Tutorial Team, Advanced Git: Understanding Git Collaboration & Workflows, 2nd ed. (New York: Razeware LLC, 2021). 4) T. Carter, Mastering Git for Version Control: Optimize Collaborative Workflows and Conflict Resolution with Git's Most Advanced Features (Independently Published, 2024). 5) Piyush Mishra, Ultimate Git and GitHub for Modern Software Development: Unlock the Power of Git and GitHub Version Control and Collaborative Coding to Seamlessly Manage and Streamline Software Projects (New Delhi: Orange Education Pvt. Ltd., 2024). 6) Douglas Cardoso, GIT Expert: Guide to Version Control Mastery (São Paulo: Douglas Cardoso Technology Publications, 2025).		
Evaluation Pattern:		
Total Marks: 50		
Internal Continuous Evaluation (20 Marks): • CCE - I: 10 Marks: Activity Based • CCE - II: 10 Marks: Activity Based • CCE – III: 10 Marks: MCQ Based Note: Best two scores will be considered for final mark entry.		End Semester Examination (30 Marks): • Question -1: Solve the following questions (Five questions of 2 Marks) • Question -2: Attempt any two questions (Three questions of 10 Marks) • Question -3: Attempt any four questions (Five questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester - VI
	Type : Major P VII	Marks: 50
	Credits : 2	From: A. Y. 2026-27
Name of the Course and Code: Practical VII (BDSP 366)		
Course Objectives: Students should...		
1) Provide students with a comprehensive understanding of data mining. 2) Equip students with practical skills to use Python and its libraries. 3) Enable students to perform data pre-processing, transformation. 4) provide hands-on experience using software tools		
Course Outcomes: Students will be:		
1) Perform data preprocessing, cleaning, transformation, and exploratory data analysis (EDA). 2) Implement and evaluate various classification algorithms. 3) Apply clustering and association rule. 4) Analyze and implement image segmentation and feature extraction methods.		
List of Experiments		Hrs
1) Demonstrate basic Python operations and time-series data handling. 2) Parse date and time values, set the date as an index, examine data types, summary statistics, missing values and perform time-based indexing and slicing. 3) Handle missing values, identify and treat outliers, remove duplicate records, convert data types and prepare the dataset for analysis. 4) Identify Trend, Seasonal, Cyclical and Irregular components using appropriate graphs and statistical techniques. 5) Decompose a time-series dataset using both models and compare the trend, seasonal, and residual components. 6) Handle missing values before decomposition, visualize and interpret the trend, seasonal and residual components. 7) Simple Moving Average (SMA), Weighted Moving Average (WMA), and Exponential Smoothing to a dataset and visualize the smoothed series. 8) Prepare the dataset, build an ARIMA model, generate forecasts, and visualize actual versus predicted values. 9) Train both models on the same dataset and generate future predictions. 10) Perform stock-market, weather, or sales/demand analysis using Pandas, Matplotlib,		60

Seaborn and identify important patterns and trends. 11) Model Evaluation using RMSE, MSE, MAE with Graphical Comparison 12) Install Git and Configure User Details and Initialize a Git Repository and Make the First Commit 13) Track File Changes Using Git (git status, git diff) and Undo Changes: git reset, git revert, git checkout 14) Create and Switch Between Branches, Merge Branches and Resolve Conflicts 15) Fork a Repository and Create a Pull Request and collaborate with GitHub Issues and Discussions	
Reference Books:- 1) Han, J., Kamber, M., & Pei, J. <i>Data Mining: Concepts and Techniques</i>. San Francisco: Morgan Kaufmann, 2011. 2) McKinney, W. <i>Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython</i>. Sebastopol: O'Reilly Media, 2017. 3) Raschka, S. <i>Python Machine Learning</i>. Birmingham: Packt Publishing, 2015. 4) Tan, P.-N., Steinbach, M., & Kumar, V. <i>Introduction to Data Mining</i>. Boston: Addison-Wesley, 2005. 5) Leskovec, J., Rajaraman, A., & Ullman, J. D. <i>Mining of Massive Datasets Using Python</i>. Cambridge: Cambridge University Press, 2020. 6) Patel, R., & Mehta, R. <i>Data Mining Using Python</i>. New Delhi: BPB Publications, 2021. 7) McKinney, Wes. <i>Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter</i>. 3rd ed. Sebastopol, CA: O'Reilly Media, 2022. 8) VanderPlas, Jake. <i>Python Data Science Handbook: Essential Tools for Working with Data</i>. 2nd ed. Sebastopol, CA: O'Reilly Media, 2022.	
Evaluation Pattern:	
Total Marks: 50	
Journal, Students' Performance Viva, Project (20 Marks): • Journal: 10 Marks • Students' Performance: 05 Marks • Viva: 5 Marks	Practical Exam Paper (30 Marks): • Section I: Attempt any two questions (Three questions of 10 Marks) • Section II: Attempt any two questions (Two questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology		
	Yashavantrao Chavan Institute of Science, Satara		
	Board of Studies in Statistics (Data Science)		
	Programme: B. Sc.	Semester - VI	
	Type: Major XVI (Elective)	Marks: 50	
	Credits: 2	From: A. Y. 2026-27	
Name of the Course and Code: Data Engineering (BDST 364)			
Course Objectives: Students should be able to...			
1) Develop the skills of managing the data with respect to knowledge generation. 2) Provide the ability to design the data engineering process. 3) Propose the data reliability models. 4) Define how to use machine learning models			
Course Outcomes: Students will be ...			
1) Building storage system with appropriate data technologies. 2) Designing data pipelines and data flow. 3) Processing the data infrastructure demonstrate use of advanced foss computing environments for big health care data. 4) Training and measuring the serving infrastructure for machine learning models			
Module	Title and Contents	Hrs	
Module 1:	Data Engineering Concepts 1.1 Introduction and Fundamentals of Data Engineering, Overview Of Data Engineering: Role, Importance and Workflows, Data Engineering Life Cycle. 1.2 Data Types and Structure, Structured, semi-structured, and unstructured data, Schema design considerations, 1.3 Building and Operationalizing Storage Systems, Cloud SQL: Use cases and operations, Cloud Spanner: Global RDBMS for scalable workloads, Cloud Bigtable: NoSQL wide-column database, 1.4 Cloud Firestore: Document database, BigQuery: Serverless data warehouse, Cloud Memorystore: Redis and Memcached, Cloud Storage: Object storage system	08	
Module 2:	Designing Data Pipelines 2.1 Overview of Data Pipelines, Lifecycle and components of a pipeline		

	2.2 Google Cloud Platform (GCP) Pipeline Components, Cloud Pub/Sub, Dataflow, Big Query, Cloud Functions, Integration and orchestration 2.3 Migrating Hadoop and Spark to GCP, GCP alternatives to Hadoop ecosystem, Migration best practices and tools 2.4 Designing a Data Processing Solution, Infrastructure design for distributed processing, Building scalable and fault-tolerant pipelines, Migrating traditional data warehouses to BigQuery	07
Module 3:	Designing for Security and Compliance 3.1 Identity and Access Management (IAM), Principles of IAM in GCP, Using IAM with storage and processing services 3.2 Data Security and Privacy, Encryption, key management, and secure access, Data Loss Prevention API: Classification and redaction, Privacy-preserving analytics 3.3 Legal and Regulatory Compliance, Understanding GDPR, HIPAA, and other regulations, Ensuring compliance in cloud data projects	07
Module 4:	Data Operations for Flexibility and Portability 4.1 Cataloging and Data Discovery, Google Data Catalog overview, Metadata management 4.2 Data Preprocessing and Visualization, Data wrangling with Dataprep, Visualization using Data Studio 4.3 Exploring and Analyzing Data, Cloud Datalab for data exploration, Use cases and example notebooks	08
Reference Books:- 1) Steven L. Brunton and J. Nathan Kutz, Data-Driven Science and Engineering, Cambridge: Cambridge University Press, 2019. 2) M. Housley and A. Verner, Fundamentals of Data Engineering, Cambridge: Cambridge University Press, 2019. 3) Martin Kleppmann, Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems, Cambridge: Cambridge University Press, 2019. 4) R. J. Schiller and D. Larochelle, Data Engineering Best Practices, Cambridge: Cambridge		

University Press, 2019.

- 5) V. Kumar, S. Chaisiri, and R. Ko, Data Security in Cloud Computing, London: The Institution of Engineering and Technology, 2020.
- 6) D. Sullivan, Professional Data Engineer, Indianapolis: Sybex–Wiley, 2020.
- 7) Vlad Riscutia, Data Engineering on Azure, Shelter Island, NY: Manning Publications, 2021.

Evaluation Pattern:

Total Marks: 50

Internal Continuous Evaluation (20 Marks):

- CCE - I: 10 Marks: Activity Based
- CCE - II: 10 Marks: Activity Based
- CCE – III: 10 Marks: MCQ Based

Note: Best two scores will be considered for final mark entry.

End Semester Examination (30 Marks):

- Question -1: Solve the following questions (Five questions of 2 Marks)
- Question -2: Attempt any two questions (Three questions of 10 Marks)
- Question -3: Attempt any four questions (Five questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester - VI
	Type : Major P VIII (Elective)	Marks: 50
	Credits : 2	From: A. Y. 2026-27
Name of the Course and Code: Practical VIII (BDSP 367)		
Course Objectives: Students should be ... 1) Develop the skills of managing the data with respect to knowledge generation. 2) Provide the ability to design the data engineering process. 3) Propose the data reliability models. 4) Define how to use machine learning models		
Course Outcomes: Students will be ... 1) Building the storage system with appropriate data technologies 2) Designing the data pipelines and data flow 3) Processing the data infrastructure demonstrate use of advanced Foss Computing environments for big health care data. 4) Training and measuring the serving infrastructure for machine learning models		
List of Experiments		Hrs
1) Demonstration of Cloud Spanner. 2) Practical on unmanaged databases. 3) Exercise on GCP pipeline components. 4) Demonstration of migrating Hadoop and Spark to GCP. 5) Practical on designing infrastructure for distributed data processing. 6) Data Engineering Lifecycle Case Study (End-to-End Workflow Design) 7) Structured vs Semi-Structured Data Handling Case Study 8) Schema Design for E-commerce Database Case Study 9) Cloud Storage Selection Case Study (SQL vs NoSQL vs BigQuery) 10) Building a Data Lake using Cloud Storage Case Study 11) Real-Time Data Pipeline using Pub/Sub Case Study 12) Migrating Hadoop-Based System to GCP Case Study 13) Designing Scalable and Fault-Tolerant Pipeline Case Study		60

14) IAM Role-Based Access Control Implementation Case Study		
15) GDPR/Compliance Implementation in Data Pipeline Case Study		
Reference Books:- 1) Steven L. Brunton and J. Nathan Kutz, Data-Driven Science and Engineering (Cambridge: Cambridge University Press, 2019). 2) Martin Kleppmann, Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems (Cambridge: Cambridge University Press, 2019). 3) R. J. Schiller and D. Larochelle, Data Engineering Best Practices (Cambridge: Cambridge University Press, 2019). 4) V. Kumar, S. Chaisiri, and R. Ko, Data Security in Cloud Computing (London: The Institution of Engineering and Technology, 2020). 5) Dan Sullivan, Professional Data Engineer (Indianapolis: Sybex–Wiley, 2020). 6) Vlad Riscutia, Data Engineering on Azure (Shelter Island, NY: Manning Publications, 2021). 7) M. Housley and A. Verner, Fundamentals of Data Engineering (Shelter Island, NY: Manning Publications, 2021).		
Evaluation Pattern:		
Total Marks: 50		
Journal, Students' Performance Viva, Project (20 Marks): • Journal: 10 Marks • Students' Performance: 05 Marks • Viva: 5 Marks		Practical Exam Paper (30 Marks): • Section I: Attempt any two questions (Three questions of 10 Marks) • Section II: Attempt any two questions (Two questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B. Sc.	Semester - VI
	Type: Major XVI (Elective)	Marks: 50
	Credits: 2	From: A. Y. 2026-27
	Name of the Course and Code: Data Governance Methodology (BDST 365)	
Course Objectives: Students should be ... 1)Develop the skills of managing the data with respect to knowledge generation. 2) Provide the ability to design the data engineering process. 3) Propose the data reliability models. 4) Define how to use machine learning models		
Course Outcomes: Students will be ... 1) Building the storage system with appropriate data technologies. 2) Designing the data pipelines and data flow. 3) Processing the data infrastructure. 4) Training and measuring the serving infrastructure for machine learning models		
Module	Title and Contents	Hrs
Module 1:	Foundations of Data Governance and Its Strategic Business Value 1.1 Understanding Data Governance, Definition of Data Governance, Importance and relevance in modern organizations 1.2 Business Value of Data Governance, Strategic advantages, Enhancing decision-making and compliance, Risk mitigation and cost efficiency 1.3 Developing and Preparing for Governance, Steps in developing a data governance framework, Readiness assessment for implementation	08
Module 2:	Implementing Data Governance People, Processes, and the Data Life Cycle 2.1 Core Components, The Enterprise Data Dictionary, People: Roles, responsibilities, and wearing multiple hats 2.2 Process Considerations, Adapting to diverse organizational needs, Common issues and strategic approaches, Integrating people and process for success 2.3 Phases of the data life cycle, Data life cycle management principles, Operationalizing data governance effectively.	07
Module 3:	Foundations of Data Governance and Its Strategic Business Value	

	3.1 Fostering a Data Culture, establishing a culture of data responsibility, Business benefits of data governance culture, Aligning intention, training. 3.2 Maintaining Agility and Compliance, Keeping governance agile and scalable, Legal and security collaboration, Incident handling and transparency mechanisms 3.3 Addressing Governance Challenges, Risk response strategies, Dealing with data governance obstacles.	07
Module 4:	Data Protection Through Planning Advanced Techniques and Monitoring 4.1 Planning for Data Protection, Fundamentals of data protection, Data protection in cloud and hybrid environments, Physical security of data assets 4.2 Advanced Protection Techniques, Preventing data exfiltration, Identity and access management, Agile practices in data protection, Industry best practices 4.3 Monitoring and Evaluation, Identifying who should monitor data, Components of an effective monitoring system.	08

Reference Books:-

- 1) Steven L. Brunton and J. Nathan Kutz, Data-Driven Science and Engineering (Cambridge: Cambridge University Press, 2019).
- 2) M. Housley and A. Verner, Fundamentals of Data Engineering (Cambridge: Cambridge University Press, 2019).
- 3) Martin Kleppmann, Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems (Cambridge: Cambridge University Press, 2019).
- 4) R. J. Schiller and D. Larochelle, Data Engineering Best Practices (Cambridge: Cambridge University Press, 2019).
- 5) V. Kumar, S. Chaisiri, and R. Ko, Data Security in Cloud Computing (London: The Institution of Engineering and Technology, 2020).
- 6) D. Sullivan, Professional Data Engineer (Indianapolis: Sybex–Wiley, 2020).
- 7) Vlad Riscutia, Data Engineering on Azure (Shelter Island, NY: Manning Publications, 2021).

Evaluation Pattern:

Total Marks: 50

Internal Continuous Evaluation (20 Marks): • CCE - I: 10 Marks: Activity Based • CCE - II: 10 Marks: Activity Based • CCE – III: 10 Marks: MCQ Based Note: Best two scores will be considered for final mark entry.	End Semester Examination (30 Marks): • Question -1: Solve the following questions (Five questions of 2 Marks) • Question -2: Attempt any two questions (Three questions of 10 Marks) • Question -3: Attempt any four questions (Five questions of 5 Marks) Note: Conversion of 50 marks of ESE evaluation to 30Marks
--	---

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester - VI
	Type : Major P VIII (Elective)	Marks: 50
	Credits : 2	From: A. Y. 2026-27
Name of the Course and Code: Practical VIII (BDSP 367)		
Course Objectives: Students should be		
1) Conceptualize the data governance. 2) Acquire the legal framework and essential compliances for data governance. 3) Understanding the concern of privacy and security in data governance modules. 4) Comprehend the ethical issues involved in data governance		
Course Outcomes: Students will be ...		
1) Demonstrate a comprehensive understanding of data governance, including its importance and benefits for organizations. 2) Analyze real-world examples of data governance in action and evaluate their impact on business value. 3) Apply data governance principles and tools, such as the enterprise dictionary and policy management, to support effective data management practices. 4) Implement data governance practices throughout the data life cycle and effectively operationalize data governance within an organization.		
List of Experiments		Hrs
1) Demonstration of Data Governance. 2) Practical on Business Value of Data Governance. 3) Hands on Enterprise Dictionary. 4) Practical on The Process- Diverse Companies. 5) Practical on Interplay with Legal and Security 6) Hands on on Incident Handling 7) Study of Data Governance Concepts and Organizational Importance 8) Analysis of Business Value and Strategic Benefits of Data Governance 9) Designing a Data Governance Framework for an Organization 10) Case Study on Big Data Impact and Data-Driven Decision Making 11) Development of Enterprise Data Dictionary 12) Study of Roles and Responsibilities in Data Governance (People Aspect)		60

13) Analysis of Data Life Cycle and Governance Implementation 14) Implementation of Data Governance Processes in an Organization 15) Planning and Implementation of Data Protection and Monitoring Techniques	
Reference Books:- 1) Steven L. Brunton and J. Nathan Kutz, Data-Driven Science and Engineering (Cambridge: Cambridge University Press, 2019). 2) Martin Kleppmann, Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems (Cambridge: Cambridge University Press, 2019). 3) R. J. Schiller and D. Larochelle, Data Engineering Best Practices (Cambridge: Cambridge University Press, 2019). 4) V. Kumar, S. Chaisiri, and R. Ko, Data Security in Cloud Computing (London: The Institution of Engineering and Technology, 2020). 5) Dan Sullivan, Professional Data Engineer (Indianapolis: Sybex–Wiley, 2020). 6) Vlad Riscutia, Data Engineering on Azure (Shelter Island, NY: Manning Publications, 2021). 7) M. Housley and A. Verner, Fundamentals of Data Engineering (Shelter Island, NY: Manning Publications, 2021).	
Evaluation Pattern:	
Total Marks: 50	
Journal, Students' Performance Viva, Project (20 Marks): • Journal: 10 Marks • Students' Performance: 05 Marks - Viva: 5 Marks	Practical Exam Paper (30 Marks): • Section I: Attempt any two questions (Three questions of 10 Marks) • Section II: Attempt any two questions (Two questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc.	Semester - VI
	Type: VSC IV	Marks: 50
	Credits: 2	From: A. Y. 2026-27
	Name of the Course and Code: Data Visualization Using Power BI (BDSPVSC IV)	
Course Objectives: Students should be ... 1) Provide a comprehensive understanding of Power BI tools and their applications in business intelligence. 2) Enable students to effectively clean, transform, and model data for analysis and reporting. 3) Teach the creation of visually appealing and interactive dashboards and reports. 4) Familiarize students with advanced features, including DAX, dataflows and Power BI integration.		
Course Outcomes: Students will be ... 1) Connect to various data sources, clean, and transform data using power query. 2) Develop and optimize data models, leveraging dax for calculations and analysis. 3) Create professional-grade reports and dashboards with advanced interactivity. 4) Deploy and manage power bi solutions, ensuring security and collaboration in real-world scenarios.		
List of Experiments		Hrs
1. Download and install Power BI Desktop and explore the interface. 2. Import data from Excel, SQL Server, and web data sources. 3. Creation of charts: Pie charts, bar chart, histogram 4. Perform operations like removing duplicates, splitting columns, and handling null value using power query. 5. Build and manage relationships between tables in Power BI. 6. Create a star schema data model and optimize for analytics. 7. Implement calculated columns for custom data fields. 8. Use basic DAX functions like SUM, AVERAGE, and COUNT. 9. Build year-over-year and month-to-date measures using DAX.		60

10. Apply themes, conditional formatting, and tooltips in report. 11. Create maps and work with location-based data 12. Combine visuals into an interactive dashboard using bookmarks. 13. Publish a Power BI report to the Power BI Service and set up scheduled refresh. 14. Implement RLS to restrict data access for specific roles. 15. View and interact with dashboards on mobile devices.	
Reference Books:- 1. Aspin, A. High Impact Data Visualization with Power View, Power Map, and Power BI. Apress, 2014. 2. Deckler, G. Learn Power BI: A Comprehensive Guide to Creating Reports and Dashboards. Packt Publishing, 2023. 3. Ferrari, A., & Russo, M. The Definitive Guide to DAX: Business Intelligence with Microsoft Excel, SQL Server Analysis Services, and Power BI. Microsoft Press, 2022. 4. Powell, B. Mastering Microsoft Power BI: Expert Techniques for Effective Data Analytics and Business Intelligence. Packt Publishing, 2022. 5. Rad, R. Pro Power BI Architecture: Sharing, Security, and Deployment Options for Microsoft Power BI Solutions. Apress, 2020. 6. Raviv, G. Collect, Combine, and Transform Data Using Power Query in Excel and Power BI. Microsoft Press, 2018. 7. Seamark, P. DAX Patterns: Practical Solutions for Business Intelligence and Data Analysis. Apress, 2018	
Evaluation Pattern:	
Total Marks: 50	
Journal, Students' Performance Viva, Project (20 Marks): • Journal: 10 Marks • Students' Performance: 05 Marks - Viva: 5 Marks	Practical Exam Paper (30 Marks): • Section I: Attempt any two questions (Three questions of 10 Marks) • Section II: Attempt any two questions (Two questions of 5 Marks)

	Karmaveer Bhaurao Patil University, Satara (A State Public University Est. u/s 3(6) of MPUA 2016) Faculty of Science and Technology	
	Yashavantrao Chavan Institute of Science, Satara	
	Board of Studies in Statistics (Data Science)	
	Programme: B.Sc. III	Semester - VI
	Type :SEC -III	Marks: 50
	Credits : 2	From: A. Y. 2026-27
Name of the Course and Code: Healthcare Analytics (BDSPSEC III)		
Course Objectives: Students should be ... 1) Understand the fundamentals of healthcare data, its types, sources, and challenges 2) Learn data preprocessing techniques including data cleaning, transformation, and handling missing values 3) Develop analytical skills to extract meaningful insights from healthcare datasets 4) Gain hands-on experience in Python for data analysis and basic predictive modeling		
Course Outcomes: Students will be ... 1) Perform data preprocessing tasks such as handling missing values, outliers, and encoding using Python 2) Apply statistical and analytical techniques to analyze healthcare data and identify patterns 3) Build and evaluate basic machine learning models for healthcare prediction tasks 4) Use Python libraries to perform data analysis and generate insights for healthcare decision-making		
List of Experiments		Hrs
1) Import a healthcare dataset using Python (Pandas) 2) Perform initial data inspection using head(), tail(), info(), and describe() 3) Identify and classify variables as categorical, numerical, and binary 4) Perform data filtering, sorting, subsetting, column renaming, and data type modification 5) Handle missing values using mean, median, mode, and deletion methods 6) Identify and remove duplicate records and data inconsistencies 7) Detect outliers using IQR and Z-score methods 8) Treat outliers using appropriate techniques 9) Perform data normalization and standardization		60

10) Encode categorical variables using Label Encoding and One-Hot Encoding 11) Generate summary statistics for the dataset 12) Calculate healthcare KPIs such as mortality rate, recovery rate, and readmission rate 13) Split dataset into training and testing sets 14) Build a Logistic Regression model for classification (e.g., disease prediction) 15) Evaluate model performance using accuracy, confusion matrix, and basic evaluation metrics.	
Reference Books: - 1. Reddy, C. K., & Aggarwal, C. C. <i>Healthcare Data Analytics</i>. Boca Raton: CRC Press, 2015. 2. Belle, A., Thiagarajan, R., Soroushmehr, S. M. R., Navidi, F., Beard, D. A., & Najarian, K. <i>Big Data Analytics in Healthcare</i>. BioMed Research International, 2015. 3. Wickham, H., & Grolemund, G. <i>R for Data Science</i>. Sebastopol: O'Reilly Media, 2017. <i>(Useful for data handling concepts applicable to Python as well)</i> 4. VanderPlas, J. <i>Python Data Science Handbook</i>. Sebastopol: O'Reilly Media, 2016. 5. Géron, A. <i>Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow</i>. Sebastopol: O'Reilly Media, 2019. 6. McKinney, W. <i>Python for Data Analysis</i>. Sebastopol: O'Reilly Media, 2018. 7. Sharma, M., & Mehra, R. <i>Healthcare Analytics using Python</i>. New Delhi: BPB Publications, 2021.	
Evaluation Pattern:	
Total Marks: 50	
Journal, Students' Performance Viva, Project (20 Marks): • Journal: 10 Marks • Students' Performance: 05 Marks • Viva: 5 Marks	Practical Exam Paper (30 Marks): • Section I: Attempt any two questions (Three questions of 10 Marks) • Section II: Attempt any two questions (Two questions of 5 Marks)